

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 22, 2003, 15:32:02 ; Search time 26 Seconds

(without alignments)  
3701.175 Million cell updates/sec

Title: US-10-046-433-40

Perfect score: 1001

Sequence: 1 MAPPGHSHLSARVGRTER.....LGRSHLPPRLIMLTQCR 1001

Scoring table: OLIGO

Searched: 283224 seqs, 96134422 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 150 summaries

Database: PIR 73:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	0.8	133	2 D48776	polyprotein (E2/NS
2	8	0.8	284	2 A69536	mevalonate kinase
3	8	0.8	353	1 OPRHC	peroxidase (EC 1.1
4	8	0.8	353	2 S37495	peroxidase (EC 1.1
5	8	0.8	353	2 T46118	peroxidase - Arabi
6	8	0.8	354	2 J00457	peroxidase (EC 1.1
7	8	0.8	443	2 A71686	tolB protein precu
8	8	0.8	444	2 F97750	hypothetical prote
9	8	0.8	695	2 E72075	hypothetical prote
10	8	0.8	695	2 F86547	hypothetical prote
11	8	0.8	695	2 F81592	hypothetical prote
12	8	0.8	744	2 T39841	topoisomerase II as
13	7	0.7	73	1 DMBPD8	DNA-binding protei
14	7	0.7	75	2 T38534	small nuclear ribo
15	7	0.7	91	2 G84870	hypothetical prote
16	7	0.7	91	2 T47805	U6 snRNA-associate
17	7	0.7	100	1 B43719	urease (EC 3.5.1.5
18	7	0.7	100	1 S08478	urease (EC 3.5.1.5
19	7	0.7	100	2 B47090	urease (EC 3.5.1.5
20	7	0.7	106	2 C95075	hypothetical prote
21	7	0.7	116	2 T47435	hypothetical prote
22	7	0.7	124	2 T24876	hypothetical prote
23	7	0.7	134	2 C90490	conserved hypochet
24	7	0.7	159	2 H84224	hypothetical prote
25	7	0.7	175	2 P97734	formylmethionine d
26	7	0.7	175	2 B64396	conserved hypochet
27	7	0.7	181	2 S78336	conserved hypochet
28	7	0.7	197	2 F75498	probable phosphogl
29	7	0.7	215	2 AD0056	

30	7	0.7	215	2 A91298	phosphoglyceromuta
31	7	0.7	215	2 C86139	phosphoglyceromuta
32	7	0.7	215	2 S56619	gmbp protein - Esc
33	7	0.7	223	2 C71843	probable uroporphyr
34	7	0.7	226	2 H64672	uroporphyrinogen I
35	7	0.7	226	2 JC5966	transmembrane prot
36	7	0.7	239	2 S78180	succinate dehydrog
37	7	0.7	244	2 T41234	translation initia
38	7	0.7	247	2 T46968	flavoprotein [lipo
39	7	0.7	267	2 S40006	tryptan (EC 3.4.21
40	7	0.7	275	2 H96009	probable modificat
41	7	0.7	282	2 AG3112	hypothetical prote
42	7	0.7	282	2 AF2179	hypothetical prote
43	7	0.7	284	2 P98174	oligopeptide ABC t
44	7	0.7	292	2 A71065	probable sugar-bin
45	7	0.7	294	2 A87291	succinyl-CoA synth
46	7	0.7	294	2 T29272	hypothetical prote
47	7	0.7	300	2 E97675	succinyl-CoA synth
48	7	0.7	300	2 AD2900	succinyl-CoA synth
49	7	0.7	300	2 AR3259	succinate-CoA liga
50	7	0.7	300	2 S47764	dipeptide transpor
51	7	0.7	300	2 F91181	dipeptide transpor
52	7	0.7	300	2 B86028	dipeptide transpor
53	7	0.7	300	2 A10486	dipeptide transpor
54	7	0.7	300	2 AG0983	dipeptide transpor
55	7	0.7	303	2 B6591	CT580 hypothetical
56	7	0.7	303	2 D72031	hypothetical
57	7	0.7	306	2 UC4347	uroporphyrin-III C
58	7	0.7	310	2 AB0078	lysr-family regula
59	7	0.7	315	2 C97224	hypothetical prote
60	7	0.7	318	2 J02217	peroxidase (EC 1.1
61	7	0.7	320	2 S01774	stable inheritance
62	7	0.7	320	2 A24920	36K para protein -
63	7	0.7	327	2 F82904	conserved hypochet
64	7	0.7	336	2 A66997	ferriochrome transp
65	7	0.7	346	2 S55491	probable galactosy
66	7	0.7	350	1 ADECHR	2-dehydro-3-deoxy-
67	7	0.7	350	2 AG0593	phospho-2-dehydro-
68	7	0.7	350	2 P90726	2-dehydro-3-deoxy-
69	7	0.7	350	2 G85577	2-dehydro-3-deoxy-
70	7	0.7	360	2 A97788	microcin C7 self-I
71	7	0.7	362	2 I64128	2-dehydro-3-deoxy-
72	7	0.7	371	2 F87100	L-alanine dehydrog
73	7	0.7	371	2 T20093	hypothetical prote
74	7	0.7	386	2 A70791	probable anion tra
75	7	0.7	387	2 T12456	hypothetical prote
76	7	0.7	389	2 C82637	porphyrin biosynth
77	7	0.7	391	2 G89822	hypothetical prote
78	7	0.7	423	2 AH0334	probable long-chain
79	7	0.7	434	2 A12187	histidinol dehydro
80	7	0.7	434	2 H80131	regulator of epide
81	7	0.7	436	2 T36706	hypothetical prote
82	7	0.7	439	2 J01729	anayrin-repeat pro
83	7	0.7	450	2 A44751	carotenoid-binding
84	7	0.7	451	2 S45051	tubulin alpha-B ch
85	7	0.7	451	2 T47824	hypothetical prote
86	7	0.7	457	2 AB1104	DNA repair protein
87	7	0.7	471	2 T33787	hypothetical prote
88	7	0.7	473	2 E84853	hypothetical prote
89	7	0.7	487	1 A64472	carbamoyl-phosphat
90	7	0.7	495	2 AF1130	transcription regu
91	7	0.7	508	2 T01937	hypothetical prote
92	7	0.7	515	2 S62321	penton base protei
93	7	0.7	524	2 UB0335	nucleolar protein
94	7	0.7	530	2 T51836	high affinity nitr
95	7	0.7	530	2 B89771	hypothetical prote
96	7	0.7	533	2 A31760	Ro/SS-A complex, 6
97	7	0.7	568	2 T36704	probable dehydroge
98	7	0.7	591	2 A30098	developmental prot
99	7	0.7	603	2 T38920	methyleucetraxhyd
100	7	0.7	611	2 T38908	probable gamma-glu
101	7	0.7	618	2 D64472	carbamoyl-phosphat
102	7	0.7	629	2 T06675	hypothetical prote

phosphoglyceromuta  
phosphoglyceromuta  
gmbp protein - Esc  
probable uroporphyr  
uroporphyrinogen I  
transmembrane prot  
succinate dehydrog  
translation initia  
flavoprotein [lipo  
tryptan (EC 3.4.21  
probable modificat  
hypothetical prote  
hypothetical prote  
oligopeptide ABC t  
probable sugar-bin  
succinyl-CoA synth  
hypothetical prote  
succinyl-CoA synth  
succinate-CoA liga  
dipeptide transpor  
dipeptide transpor  
dipeptide transpor  
dipeptide transpor  
dipeptide transpor  
dipeptide transpor  
CT580 hypothetical  
hypothetical  
uroporphyrin-III C  
lysr-family regula  
hypothetical prote  
peroxidase (EC 1.1  
stable inheritance  
36K para protein -  
conserved hypochet  
ferriochrome transp  
probable galactosy  
2-dehydro-3-deoxy-  
phospho-2-dehydro-  
2-dehydro-3-deoxy-  
2-dehydro-3-deoxy-  
microcin C7 self-I  
2-dehydro-3-deoxy-  
L-alanine dehydrog  
hypothetical prote  
probable anion tra  
hypothetical prote  
porphyrin biosynth  
hypothetical prote  
probable long-chain  
histidinol dehydro  
regulator of epide  
hypothetical prote  
anayrin-repeat pro  
carotenoid-binding  
tubulin alpha-B ch  
hypothetical prote  
DNA repair protein  
hypothetical prote  
hypothetical prote  
carbamoyl-phosphat  
transcription regu  
hypothetical prote  
penton base protei  
nucleolar protein  
high affinity nitr  
hypothetical prote  
Ro/SS-A complex, 6  
probable dehydroge  
developmental prot  
methyleucetraxhyd  
probable gamma-glu  
carbamoyl-phosphat  
hypothetical prote

A;Molecule type: DNA  
A;Residues: 31-338 <GAZ>

C,Genetics: A:Gene: prxcl; w117f  
 C:Superfamily: peroxidase  
 C:Keywords: chromoprotein; glycoprotein; heme; iron; metalloprotein; oxidoreductase; pyr  
 F:1-30/Domain: signal sequence #status predicted <SIG>  
 F:31-338/Product: peroxidase C1 #status experimental <MAT>  
 F:31/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen  
 F:41-121,74-79,127-331,207-239/Disulfide bonds: #status experimental  
 F:43,87,188,216,228,244,285,298/Binding site: carbohydrate (Asn) (covalent) #status exp  
 F:68/Active site: Arg #status predicted  
 F:72,200/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 0.8%; Score 8; DB 1; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 680 VTLAGPS 687  
 DB 139 VTLAGPS 146

## RESULT 4

S37495  
 Peroxidase (EC 1.11.1.7) Cb - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 09-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 04-Mar-2000  
 C:Accession: S37495  
 R:Intapuk, C.; Takano, M.; Shimmyo, A.  
 Submitted to the EMBL Data Library, April 1993  
 A:Description: Nucleotide sequence of a new cDNA for peroxidase from Arabidopsis thaliana  
 A:Reference number: S37495  
 A:Accession: S37495  
 A:Molecule type: mRNA  
 A:Residues: 1-353 <INT>  
 A:Cross-references: EMBL:X71794; NID:9405610; PIDN:CAA50677.1; PID:9405611  
 C:Genetics: A:Gene: prx3b  
 C:Superfamily: peroxidase  
 C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
 F:41-121/Disulfide bonds: #status predicted  
 F:68/Active site: Arg #status predicted  
 F:72,200/Binding site: heme iron (His) (axial ligands) #status predicted  
 F:74-79/Disulfide bonds: #status predicted  
 F:127-331/Disulfide bonds: #status predicted  
 F:207-239/Disulfide bonds: #status predicted

Query Match 0.8%; Score 8; DB 2; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 680 VTLAGPS 687  
 DB 139 VTLAGPS 146

## RESULT 5

T46118  
 peroxidase - Arabidopsis thaliana  
 N:Alternate names: protein T2J13.40  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 18-Feb-2000  
 C:Accession: T46118  
 R:Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Mayer, K.  
 Submitted to the Protein Sequence Database, November 1999  
 A:Reference number: Z23023  
 A:Accession: T46118  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-353 <RIE>  
 A:Cross-references: EMBL:AL132967  
 A:Experimental source: cultivar Columbia; BAC clone T2J13  
 C:Genetics: A:Map position: 3

A:Introns: 77/3; 141/3; 198/1  
 A:Note: T2J13.40  
 C:Superfamily: peroxidase

Query Match 0.8%; Score 8; DB 2; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 680 VTLAGPS 687  
 DB 139 VTLAGPS 146

## RESULT 6

JU0457  
 peroxidase (EC 1.11.1.7) C - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 04-Mar-2000  
 C:Accession: JU0457; T46119  
 R:Intapuk, C.; Higashimura, N.; Yamamoto, K.; Okada, N.; Shimmyo, A.; Takano, M.  
 Gene 98, 237-241, 1991  
 A:Title: Nucleotide sequences of two genomic DNAs encoding peroxidase of Arabidopsis the  
 A:Reference number: JU0457; MUID:9120671; PMID:2016063  
 A:Accession: JU0457  
 A:Molecule type: DNA  
 A:Residues: 1-354 <INT>  
 A:Cross-references: GB:MS8380; NID:9166826; PIDN:AAA32849.1; PID:9166827  
 R:Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Mayer, K.  
 Submitted to the Protein Sequence Database, November 1999  
 A:Reference number: Z23023  
 A:Accession: T46119  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-354 <RIE>  
 A:Cross-references: EMBL:AL132967  
 A:Experimental source: cultivar Columbia; BAC clone T2J13  
 C:Genetics: A:Gene: prxCa; T2J13.50  
 A:Map position: 3  
 A:Introns: 78/3; 142/3; 199/1

C:Superfamily: peroxidase  
 C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
 F:42-122/Disulfide bonds: #status predicted  
 F:69/Active site: Arg #status predicted  
 F:73,201/Binding site: heme iron (His) (axial ligands) #status predicted  
 F:75-80/Disulfide bonds: #status predicted  
 F:128-332/Disulfide bonds: #status predicted  
 F:208-240/Disulfide bonds: #status predicted

Query Match 0.8%; Score 8; DB 2; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 680 VTLAGPS 687  
 DB 140 VTLAGPS 147

## RESULT 7

A71686  
 cold protein precursor (COLB) RP302 - Rickettsia prowazekii  
 C:Species: Rickettsia prowazekii  
 C>Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
 C:Accession: A71686  
 R:Anderson, S.G.E.; Zomrodipour, A.; Anderson, J.O.; Sichert-Ponten, T.; Alsmark, U.  
 Nature 396, 133-140, 1998  
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
 A:Reference number: A71686; MUID:99039499; PMID:9823893  
 A:Accession: A71686  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-443 <AND>  
 A:Cross-references: GB:AJ235271; GB:AJ235269; NID:93868717; PIDN:CAA14763.1; PID:93868686

A:Experimental source: strain Madrid E  
C:Gene: colB, RP302

Query Match 0.8%; Score 8; DB 2; Length 443;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 PRGDYIAF 168  
DB 340 PRGDYIAF 347

#### RESULT 8

colB protein precursor [imported] - Rickettsia conorii (strain Malish 7)  
C:Species: Rickettsia conorii  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 30-Sep-2001

R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Rickettsia conorii

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: F97750  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-444 <KUR>

A:Cross-references: GB:AE006914; PIDN:AAL02944.1; PID:G15619474; GSPDB:GN00173

C:Gene: colB

Query Match 0.8%; Score 8; DB 2; Length 444;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 PRGDYIAF 168  
DB 341 PRGDYIAF 348

#### RESULT 9

hypothetical protein - Chlamydia pneumoniae (strain CWL029)  
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 05-May-2000

R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.; Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MUID:99206606; PMID:10192388

A:Accession: E72075  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-695 <ARN>

A:Cross-references: GB:AE001630; GB:AE001363; NID:94376740; PIDN:AAD18600.1; PID:9437674

C:Gene: CP0458

C:Superfamily: Chlamydia hypothetical protein CP0462

Query Match 0.8%; Score 8; DB 2; Length 695;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 868 VSSCVAGI 875  
DB 67 VSSCVAGI 74

#### RESULT 10

hypothetical protein CP0458 [imported] - Chlamydia pneumoniae (strain J138)  
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001

C:Accession: F86547  
R:Shirai, M.; Hirakawa, H.; Kinoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.  
A:Reference number: A86491; MUID:20330349; PMID:10871362

A:Accession: F86547  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-695 <STO>

A:Cross-references: GB:BA000008; NID:98978828; PIDN:BA98664.1; GSPDB:GN00142

A:Experimental source: strain J138

C:Gene: CP10458

C:Superfamily: Chlamydia hypothetical protein CP0462

Query Match 0.8%; Score 8; DB 2; Length 695;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

#### RESULT 11

hypothetical protein CP0294 [imported] - Chlamydia pneumoniae (strain AR39)  
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-May-2000

R:Read, T.D.; Brundham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: F81592  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-695 <REA>

A:Cross-references: GB:AE002191; GB:AE002161; NID:G7189216; PIDN:AAF38151.1; PID:G718922

A:Experimental source: strain AR39, HL cells

C:Gene: CP0294

C:Superfamily: Chlamydia hypothetical protein CP0462

Query Match 0.8%; Score 8; DB 2; Length 695;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

#### RESULT 12

topoisomerase II associated protein patI homolog - fission yeast (Schizosaccharomyces pombe)

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 07-Dec-1999

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Devlin, K.; Churcher, C.M. submitted to the EMBL Data Library, February 1998

A:Reference number: Z21884  
A:Accession: T39841  
A:Status: preliminary; translated from GB/EMBL/DD8J

A:Molecule type: DNA

A:Residues: 1-744 <LYN>

A:Cross-references: EMBL:AL021839; PIDN:CAA17064.1; GSPDB:GN00067; SPDB:SPBC1967.10C

A:Experimental source: strain 972h-; cosmid c19G7

C:Gene: SPDB:SPBC1967.10C

A:Map position: 2

A;introns: 26/2; 60/1; 656/2; 713/2

Query Match 0.8%; Score 8; DB 2; Length 744;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 183 LKOSGTVA 190  
|||||  
Db 642 LKOSGTVA 649

## RESULT 13

DNEPD8  
DNA-binding protein Ner - phage D108

C;Species: phage D108  
C;Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 22-Jun-1999  
C;Accession: A24680; S07931

R;Tollas, P.P.; Dubow, M.S.

EMBO J. 4, 3031-3037, 1985.

A;Title: The cloning and characterization of the bacteriophage D108 regulatory DNA-binding

A;Reference number: A24680; MUID:86055744; PMID:2998774

A;Accession: A24680

A;Molecule type: DNA

A;Residues: 1-73 <TO>

A;Cross-references: GB:M26291; NID:g16194; PIDN:AAA2206.1; PID:g16195

R;Mizunuchi, M.; Weisberg, R.A.; Mizunuchi, K.

Nucleic Acids Res. 14, 3813-3825, 1986

A;Title: DNA sequence of the control region of phage D108: the N-terminal amino acid seq

A;Reference number: S07370; MUID:86232621; PMID:3012481

A;Accession: S07931

A;Molecule type: DNA

A;Residues: 1-73 <MI>

A;Cross-references: EMBL:X03847; NID:g15439; PIDN:CAA27475.1; PID:g15441

C;Genetics:

A;Gene: ner

C;Superfamily: phage D108 DNA-binding protein

C;Keywords: DNA binding

Query Match 0.7%; Score 7; DB 1; Length 73;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 982 LGRSNHL 988  
|||||  
Db 31 LGRSNHL 37

## RESULT 14

small nuclear ribonucleoprotein, F-like - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C;Accession: J38534

R;Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, September 1997

A;Reference number: Z21748

A;Accession: J38534

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-75 <OLI>

A;Cross-references: EMBL:Z99165; PIDN:GAB54975.1; GSPDB:GN00066; SPDB:SPAC2F3.17C

A;Experimental source: strain 972h-; cosmid c2F3

C;Genetics:

A;Gene: SPDB:SPAC2F3.17C

A;Map position: 1

A;introns: 29/1; 67/1

Query Match 0.7%; Score 7; DB 2; Length 75;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 63 RGNVLY 69

## RESULT 15

hypothetical protein At2g43810 (imported) - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C;Accession: G84870

R;Lin, X.; Kaul, S.; Roundley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A64420; MUID:20083487; PMID:10617157

A;Accession: G84870

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-91 <STO>

A;Cross-references: GB:A6002093; NID:g2281089; PIDN:AA64025.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g43810

A;Map position: 2

Query Match 0.7%; Score 7; DB 2; Length 91;

Best Local Similarity 100.0%; Pred. No. 42;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 237 RGNVLY 243  
|||||  
Db 73 RGNVLY 79

## RESULT 16

U6 snRNA-associated Sm-like protein - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C;Accession: T47805

R;P'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.

submitted to the Protein Sequence Database, February 2000

A;Reference number: Z24477

A;Accession: T47805

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-91 <DAN>

A;Cross-references: EMBL:AL138647

A;Experimental source: cultivar Columbia; BAC clone F24G16

C;Genetics:

A;Map position: 3

A;introns: 39/1; 77/1

A;Note: F24G16.80

Query Match 0.7%; Score 7; DB 2; Length 91;

Best Local Similarity 100.0%; Pred. No. 42;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 237 RGNVLY 243  
|||||  
Db 73 RGNVLY 79

## RESULT 17

urease (EC 3.5.1.5) 11K chain - Proteus mirabilis

C;Species: Proteus mirabilis

C;Date: 03-Mar-1993 #sequence\_revision 02-Dec-1994 #text\_change 18-Jun-1999

C;Accession: B43719

R;Jones, B.D.; Mobley, H.L.T.

J. Bacteriol. 171, 6414-6422, 1989

A;Title: Proteus mirabilis urease: nucleotide sequence determination and comparison with

A:Reference number: A43719; MUID:90078080; PMID:2667233  
 A:Accession: B43719  
 A:Molecule type: DNA  
 A:Residues: 1-100 <JUN>  
 A:Cross-references: GB:M1834; NID:g150914; PIDN:AAA25667.1; PID:g150916  
 C:Genetics:  
 A:Gene: urea  
 A:Function:  
 A:Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide and two  
 C:Superfamily: urease 11k chain; urease 11k chain homology  
 C:Keywords: heterotrimer; hydrolase; metalloprotein  
 F:1-100/Domain: urease 11k chain homology <U11>

Query Match 0.7%; Score 7; DB 1; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 960 SCAIMEG 966  
 DB 40 SCAIMEG 46

RESULT 18  
 S08478  
 urease (EC 3.5.1.5) 11k chain - Proteus vulgaris  
 N:Alternate names: urease gamma chain  
 C:Species: Proteus vulgaris  
 C:Date: 29-Jan-1993 #sequence\_revision 02-Dec-1994 #text\_change 20-Apr-2001  
 C:Accession: S08478  
 R:Moerdorf, G.; Kaltwasser, H.  
 FEMS Microbiol. Lett. 66, 67-74, 1991  
 A:Title: Cloning of the genes encoding urease from Proteus vulgaris and sequencing of th  
 A:Reference number: S08478; MUID:92038908; PMID:1936338  
 A:Accession: S08478  
 A:Molecule type: DNA  
 A:Residues: 1-100 <MOE>  
 A:Cross-references: EMBL:X51816; NID:g45933; PIDN:CAA36113.1; PID:g45934  
 C:Genetics:  
 A:Gene: urea  
 C:Function:  
 A:Description: trimer of heterotrimeric subunits consisting of 11k (gamma), 12k (beta), and  
 C:Superfamily: urease 11k chain; urease 11k chain homology  
 C:Keywords: heterotrimer; hydrolase; metalloprotein  
 F:1-100/Domain: urease 11k chain homology <U11>

Query Match 0.7%; Score 7; DB 1; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 960 SCAIMEG 966  
 DB 40 SCAIMEG 46

RESULT 19  
 B47090  
 urease (EC 3.5.1.5) chain A - Escherichia coli  
 C:Species: Escherichia coli  
 C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 22-Jun-1999  
 C:Accession: B47090  
 R:D'Orazio, S.E.; Collins, C.M.  
 J. Bacteriol. 175, 1860-1864, 1993  
 A:Title: Characterization of a plasmid-encoded urease gene cluster found in members of t  
 A:Reference number: A47090; MUID:93194816; PMID:8443894  
 A:Accession: B47090  
 A:Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-100 <D10>  
 A:Cross-references: GB:I03307; NID:g148150; PIDN:AAA24445.1; PID:g148152  
 A:Note: sequence extracted from NCBI backbone (NCBIN:127840; NCBI:P:127842)  
 C:Genetics:

A:Gene: Urea  
 C:Superfamily: urease 11k chain; urease 11k chain homology  
 C:Keywords: hydrolase  
 F:1-100/Domain: urease 11k chain homology <U11>

Query Match 0.7%; Score 7; DB 2; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 960 SCAIMEG 966  
 DB 40 SCAIMEG 46

RESULT 20  
 C95075  
 hypothetical protein SP0650 [imputed] - Streptococcus pneumoniae (strain TIGR4)  
 C:Species: Streptococcus pneumoniae  
 C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001  
 C:Accession: C95075  
 R:Uetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,  
 nson, T.; Hickey, E.K.; Holt, I.E.  
 Science 293, 498-506, 2001  
 A:Authors: Lofcus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
 A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.  
 A:Reference number: A95000; MUID:21357209; PMID:11463916  
 A:Accession: C95075  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-106 <KIR>  
 A:Cross-references: GB:AB005672; PIDN:AAK74796.1; PID:g14972123; GSPDB:GN00164; TIGR:SP4  
 A:Experimental source: strain TIGR4  
 C:Genetics:  
 A:Gene: SP0650

Query Match 0.7%; Score 7; DB 2; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 470 MILTLV 476  
 DB 1 MILTLV 7

RESULT 21  
 T47435  
 hypothetical protein T22K7.130 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 05-May-2000  
 C:Accession: T47435  
 R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: Z24459  
 A:Accession: T47435  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-116 <RIE>  
 A:Cross-references: EMBL:AL138641  
 A:Experimental source: cultivar Columbia; BAC clone T22K7  
 C:Genetics:  
 A:Map position: 3  
 A:Note: T22K7.130  
 C:Superfamily: Arabidopsis thaliana hypothetical protein T22K7.130

Query Match 0.7%; Score 7; DB 2; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 ETTVLSG 440  
 DB 34 ETTVLSG 40

## RESULT 22

T24876  
hypothetical protein T13F2.2 - *Caenorhabditis elegans*

C/Species: *Caenorhabditis elegans*  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 03-Jun-2000  
C/Accession: T24876

R/Swiftburne, J.

submitted to the EMBL Data Library, October 1996

A/Reference number: Z19947

A/Accession: T24876

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-124 <WTL>

A/Cross-references: EMBL:Z81122; PIDN:CA803353.1; GSPDB:GN00022; CESP:T13F2.2

A/Experimental source: clone T13F2

C/Genetics:

A/Gene: CESP:T13F2.2

A/Map position: 4

A/Intons: 27/3; 61/3; 91/3

C/Superfamily: *Caenorhabditis elegans* hypothetical protein T13F2.2

## Query Match

0.7%; Score 7; DB 2; Length 124;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 609 YYIDRDS 615  
|||  
DB 87 YYIDRDS 93

## RESULT 23

C90490

conserved hypothetical protein [imported] - *Sulfolobus solfataricus*

C/Species: *Sulfolobus solfataricus*

C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001

C/Accession: C90490

R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.  
arret, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A/Description: *Sulfolobus solfataricus* complete genome.

A/Reference number: A99139

A/Accession: C90490

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-134 <KUR>

A/Cross-references: GB:AE006641; NID:gl3816489; PIDN:AAK43178.1; GSPDB:GN00155

C/Genetics:

A/Gene: SSO3078

## Query Match

0.7%; Score 7; DB 2; Length 134;  
Best Local Similarity 100.0%; Pred. No. 58;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 714 CTDNVT 720  
|||  
DB 76 CTDNVT 82

## RESULT 24

H84224

hypothetical protein Vng0670h [imported] - *Halobacterium* sp. NRC-1

C/Species: *Halobacterium* sp. NRC-1

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C/Accession: H84224

R/Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Leaky, S.  
; Lettner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo

ung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li

A/Title: Genome sequence of *Halobacterium* species NRC-1.

A/Reference number: A84160; MUID:20504483; PMID:11016550

A/Accession: H84224  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-159 <STO>  
A/Cross-references: GB:AE004437; NID:gl0580256; PIDN:AA919164.1; GSPDB:GN00138  
C/Genetics:  
A/Gene: VNG0670H

## Query Match

0.7%; Score 7; DB 2; Length 159;  
Best Local Similarity 100.0%; Pred. No. 66;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LLMAGT 34  
|||  
DB 124 LLMAGT 130

## RESULT 25

E97734

formylmethionine deformylase (EC 3.5.1.31) - *Rickettsia conorii* (strain Malish 7)

C/Species: *Rickettsia conorii*

C/Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 22-Oct-2001

C/Accession: E97734

R/Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barde, V.; Samson, D.; Ro

science 293, 2093-2098, 2001

A/Title: Mechanisms of Evolution in *Rickettsia conorii* and *Rickettsia prowazekii*.

A/Reference number: A97700; MUID:21442074; PMID:11557893

A/Accession: E97734

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-175 <KUR>

A/Cross-references: GB:AE006914; PIDN:AL02816.1; PID:gl5619334; GSPDB:GN00173

C/Genetics:

A/Gene: defl

C/Superfamily: polypeptide deformylase

C/Keywords: hydrolase

Query Match  
0.7%; Score 7; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 892 ISLPEOR 898  
|||  
DB 100 ISLPEOR 106

## RESULT 26

E64396

hypothetical protein MJ0773 - *Methanococcus jannaschii*

C/Species: *Methanococcus jannaschii*

C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C/Accession: E64396

R/Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

; Reish, C.T.; Overbeek, R.; Kirkness, E.F.; Weissbrock, K.G.; Merrick, J.M.; Glodok, A.;

Iron, J.D.; Sadow, P.W.; Hama, M.C.; Cotton, M.D.; Roberts, K.M.; Huret, M.A.

Science 273, 1058-1073, 1996

A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C

A/Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*

A/Reference number: A64300; MUID:96337999; PMID:8688087

A/Accession: E64396

A/Status: preliminary

A/Molecule type: nucleic acid sequence not shown; translation not shown

A/Residues: 1-175 <BDU>

A/Cross-references: GB:U67522; GB:L77117; NID:g2826315; PIDN:AA898777.1; PID:gl499593; T

C/Genetics:

A/Map position: REV94542-694015

C/Superfamily: *Methanococcus jannaschii* hypothetical protein MJ0773

Query Match  
0.7%; Score 7; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 72;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 687 SFTSKL 693



Db 68 SFTSGSL 74

## RESULT 27

S78336

conserved hypothetical protein 181 - *Odontella sinensis* chloroplastC:Species: chloroplast *Odontella sinensis*

C:Date: 17-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 20-Jun-2000

C:Accession: S78336

R:Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.

Plant Mol. Biol. Rep. 13, 336-342, 1995

A:Title: The Chloroplast Genome of a Chlorophyll a+c-containing Alga, *Odontella sinensis*

A:Reference number: S78338

A:Accession: S78336

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-181 &lt;KOW&gt;

A:Cross-references: EMBL:Z67753; NID:G1185127; PIDN:CAA91709.1; PID:G1185226

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

C:Genetics:

A:Gene: ycf4

A:genome: chloroplast

C:superfamily: conserved hypothetical protein s110226

C:keywords: chloroplast

## Query Match

Best Local Similarity 0.7%; Score 7; DB 2; Length 181;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 887 LCSGSGIS 893

Db 25 LCSGSGIS 31

## RESULT 28

F75498

conserved hypothetical protein - *Deinococcus radiodurans* (strain R1)C:Species: *Deinococcus radiodurans*

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000

C:Accession: F75498

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, H.O.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MIMD:20036896; PMID:10567266

A:Accession: F75498

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-197 &lt;WHI&gt;

A:Cross-references: GB:AE001918; GB:AE000513; NID:G6458294; PIDN:AAF10178.1; PID:G645829

C:Genetics:

A:gene: DR0598

A:Map position: 1

## Query Match

Best Local Similarity 0.7%; Score 7; DB 2; Length 197;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 458 IYTAAGA 464

Db 179 IYTAAGA 185

## RESULT 29

AD0056

probable phosphoglycerate mutase (EC 5.4.2.1) [imported] - *Yersinia pestis* (strain CO92)C:Species: *Yersinia pestis*

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AD0056

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titchell, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
 11, M.; Rutherford, K.; Simmonds, M.; Skellon, J.; Stevens, K.; Whitehead, S.; Barrall,  
 Nature 413, 523-527, 2001

A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A:Reference number: AB0001; MIMD:21470413; PMID:11586360

A:Accession: AD0056

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-215 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC89311.1; PID:G15978547; GSPDB:GN00175

C:Genetics:

A:Gene: gpmB

C:superfamily: Aquifex aeolicus phosphoglycerate mutase; phosphoglycerate mutase homolog

C:keywords: Intramolecular transference; isomerase

## Query Match

Best Local Similarity 0.7%; Score 7; DB 2; Length 215;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 722 RIBEGRS 728

Db 115 RIBEGRS 121

## RESULT 30

A91298

phosphoglyceromutase 2 [imported] - *Escherichia coli* (strain O157:H7, substrain RMD 050C:Species: *Escherichia coli*

C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 27-Nov-2001

C:Accession: A91298

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurukawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gen

A:Reference number: A99629; MIMD:21156231; PMID:11258796

A:Accession: A91298

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-215 &lt;HAY&gt;

A:Cross-references: GB:BA000007; PIDN:BA838776.1; PID:G13364831; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RMD 0509952

C:Genetics:

A:Gene: EC5353

C:superfamily: Aquifex aeolicus phosphoglycerate mutase; phosphoglycerate mutase homolog

## Query Match

Best Local Similarity 0.7%; Score 7; DB 2; Length 215;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 722 RIBEGRS 728

Db 115 RIBEGRS 121

## RESULT 31

C86139

phosphoglyceromutase 2 [imported] - *Escherichia coli* (strain O157:H7, substrain EDU933)C:Species: *Escherichia coli*

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 27-Nov-2001

C:Accession: C86139

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glanier, J.D.; Rose, D.J.; Mayhew

iller, L.; Grobeck, E.J.; Davis, N.W.; Lm, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A:Reference number: AB5480; MIMD:21074935; PMID:11206551

A:Accession: C86139

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-215 &lt;STO&gt;

A:Cross-references: GB:AE005174; NID:G12519424; PIDN:AAG59575.1; GSPDB:GN00145; UMGF:259

A:Experimental source: strain O157:H7, substrain EDU933

C:Genetics:

A:Gene: gpmB



C:Superfamily: Aquifex aeolicus phosphoglycerate mutase; phosphoglycerate mutase homolog

Query Match 0.7%; Score 7; DB 2; Length 215;

Best Local Similarity 100.0%; Pred. No. 85;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 722 RIBPGES 728  
DB 115 RIBPGES 121

#### RESULT 32

55619

gpmB protein - Escherichia coli (strain K-12)

N:Alternate names: hypothetical protein c215b

C:Species: Escherichia coli

C>Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 01-Mar-2002

C/Accession: S56619; B65255

R:Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.

Nucleic Acids Res. 23, 2105-2119, 1995

A>Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.

A/Reference number: S56314; MUID:9534362; PMID:7610040

A/Accession: S56619

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-215 <BIR>

A/Cross-references: EMBL:U14003; NID:g1263172; PIND:AAA97291.1; PID:g537235

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC

A.; Rose, D.J.; Mau, B.; Shaoh, Y.

Science 277, 1453-1462, 1997

A>Title: The complete genome sequence of Escherichia coli K-12.

A/Reference number: A64720; MUID:97426617; PMID:9278503

A/Accession: B65255

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-215 <BLAT>

A/Cross-references: GB:AE000509; GB:U00096; NID:g2367383; PIND:AACT7348.1; PID:g1790856;

A/Experimental source: strain K-12, substrain MG1655

C/Genetics:

A:Gene: gpmB

C:Superfamily: Aquifex aeolicus phosphoglycerate mutase; phosphoglycerate mutase homolog

C/Keywords: phosphotididine; phosphoprotein

F:4-185/Domain: phosphoglycerate mutase homology <PGMH>

F/9/Active site: His (phosphotididine intermediate) #struc predicted

Query Match 0.7%; Score 7; DB 2; Length 215;

Best Local Similarity 100.0%; Pred. No. 85;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 722 RIBPGES 728  
DB 115 RIBPGES 121

RESULT 33

C71843

Probable uroporphyrinogen-III synthase - Helicobacter pylori (strain J99)

C/Species: Helicobacter pylori

A/Variety: strain J99

C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999

C/Accession: C71843

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Tree, C.; Gibson, R.; Weibers, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A>Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A/Reference number: A71800; MUID:99120557; PMID:9923682

A/Accession: C71843

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-223 <ARN>

A/Cross-references: GB:AE001542; GB:AE001439; NID:g4155739; PIND:AAAD06731.1; PID:g415575

A/Experimental source: strain J99

C:Genetics:

A:Gene: hemD

Query Match 0.7%; Score 7; DB 2; Length 223;

Best Local Similarity 100.0%; Pred. No. 87;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 973 LIFTSKN 979  
DB 38 LIFTSKN 44

#### RESULT 34

H64672

uroporphyrinogen III co-synthase - Helicobacter pylori (strain 26695)

C/Species: Helicobacter pylori

C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999

C/Accession: H64672

R:Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen-

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A/Reference number: A64520; MUID:97394467; PMID:9252185

A/Accession: H64672

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-226 <TCM>

A/Cross-references: GB:AE000628; GB:AE000511; NID:g2314386; PIND:AAAD08269.1; PID:g231438

Query Match 0.7%; Score 7; DB 2; Length 226;

Best Local Similarity 100.0%; Pred. No. 88;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 973 LIFTSKN 979  
DB 38 LIFTSKN 44

#### RESULT 35

UC5966

transmembrane protein PDHAP - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe

C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Feb-2001

C/Accession: UC5966; T40882

R:Iha, H.; Takimoto, M.; Danjoh, I.; Fujiyama, A.

DNA Res. 4, 393-396, 1997

A>Title: Identification and characterization of a novel trans-membrane protein gene, pdh

A/Reference number: UC5966; MUID:98162726; PMID:9501995

A/Accession: UC5966

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-226 <IHA>

A/Cross-references: DDBJ:D88384; NID:g2911159; PIND:BA24946.1; PID:g2911160

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.

submitted to the EMBL Data Library, September 1998

A/Reference number: Z21954

A/Accession: T40882

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-226 <MOO>

A/Cross-references: EMBL:AL031764; PIND:CAA21112.1; GSPDB:GN00068; SPDB:SPCC1235.08c

A/Experimental source: strain 972h-; cosmid c1235

C/Genetics:

A:Gene: SPCC1235.08c

A/Map position: 3

C:Superfamily: Schizosaccharomyces pombe transmembrane protein PDHAP

Query Match 0.7%; Score 7; DB 2; Length 226;

Best Local Similarity 100.0%; Pred. No. 88;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 783 ELPHES 789  
 |||||  
 Db 214 ELPHES 220

## RESULT 36

succinate dehydrogenase (ubiquinone) (EC 1.3.5.1) chain 2 - *Reclinomonas americana* (ATCC 578180)  
 C/Species: *Reclinomonas americana*  
 C/Accession: ATCC 50394  
 C/Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 11-Jun-1999  
 R/Accession: S78180  
 R/Name: B.F. Burger, G. O'Kelly, C.J. Cedergren, R. Golding, G.B. Lemieux, C. Sank  
 Nature 387, 493-497, 1997  
 A/Title: An ancestral mitochondrial DNA resembling a eubacterial genome in miniature.  
 A/Reference number: S78127; MUID:97311393; PMID:9168110  
 A/Accession: S78180  
 A/Status: nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-239 <LAN>  
 A/Cross-references: EMBL:AF007261; NID:92258325; PIDN:AAD11913.1; PID:92258379  
 A/Experimental source: ATCC 50394  
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1997  
 C/Genetics:

A/Genome: mitochondrion  
 A/Superfamily: fumarate reductase iron-sulfur protein; ferredoxin 2[4Fe-4S] homology; Fe  
 C/Keywords: 2Fe-2S; 3Fe-4S; 4Fe-4S; iron-sulfur protein; metalloprotein; mitochondrion;  
 F/40-82/Domain: ferredoxin [2Fe-2S] homology <FER1>  
 F/144-226/Domain: ferredoxin 2[4Fe-4S] homology <FER2>  
 F/61, 66, 69, 81/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted  
 F/151, 154, 157, 219/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted  
 F/161, 208, 214/Binding site: 3Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 0.7%; Score 7; DB 2; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 92;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 WMKTEK 226  
 |||||  
 Db 123 WMKTEK 129

## RESULT 37

translation initiation factor eIF-6 [imported] - fission yeast (*Schizosaccharomyces pombe*)  
 C/Species: *Schizosaccharomyces pombe*  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 15-Sep-2000  
 C/Accession: T41234  
 R/Lyne, M.; Harris, D.E.; Murphey, L.D.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, January 1999  
 A/Reference number: Z21979  
 A/Accession: T41234  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-244 <LYN>  
 A/Cross-references: EMBL:AL035075; PIDN:CAA22640.1; GSPDB:GN00068; SPDB:SPCC1919.09  
 A/Experimental source: strain 972h-; cosmid c1919  
 C/Genetics:  
 A/Genome: SPDB:SPCC1919.09  
 A/Map position: 3  
 A/Intons: 12/2; 36/2  
 C/Superfamily: conserved hypothetical protein YPR016C

Query Match 0.7%; Score 7; DB 2; Length 244;  
 Best Local Similarity 100.0%; Pred. No. 94;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SLPPVK 87  
 |||||  
 Db 87 SLPPVK 93

## RESULT 38

flavoprotein [imported] - *Paracoccus denitrificans* (fragment)  
 T46968  
 C/Species: *Paracoccus denitrificans*  
 C/Date: 17-Mar-2000 #sequence\_revision 17-Mar-2000 #text\_change 21-Jul-2000  
 C/Accession: T46968  
 R/Medara, C.; Bardischewsky, F.; Friedrich, C.G.  
 J. Bacteriol. 179, 5014-5023, 1997  
 A/Title: Cloning and characterization of sulfite dehydrogenase, two c-type cytochromes,  
 sulfur oxidation.  
 A/Reference number: Z24324; MUID:97405897; PMID:9260941  
 A/Accession: T46968  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-247 <MOD>  
 A/Cross-references: EMBL:X79242; NID:92253074; PIDN:CAA55826.1; PID:92222781  
 A/Experimental source: strain GB17  
 C/Genetics:

A/Genome: soxX

Query Match 0.7%; Score 7; DB 2; Length 247;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 679 TVTLAG 685  
 |||||  
 Db 114 TVTLAG 120

## RESULT 39

trypsin (EC 3.4.21.4) precursor - African malaria mosquito  
 S4006  
 C/Species: *Anopheles gambiae* (African malaria mosquito)  
 C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 22-Jun-1999  
 C/Accession: S4006  
 R/Mueller, H.  
 submitted to the EMBL Data Library, June 1993  
 A/Reference number: S4003  
 A/Accession: S4006  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-267 <MUE>  
 A/Cross-references: EMBL:Z22930; NID:9410323; PIDN:CAA80516.1; PID:9410327  
 C/Superfamily: trypsin; trypsin homology  
 C/Keywords: hydrolase; protein digestion; serine proteinase  
 F/42-261/Domain: trypsin homology <TRY>  
 F/82,127,222/Active site: His, Asp, Ser #status predicted

Query Match 0.7%; Score 7; DB 2; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 920 AILTLV 926  
 |||||  
 Db 6 AILTLV 12

## RESULT 40

probable modification methylase protein [imported] - *Sinorhizobium meliloti* (strain 1021)  
 H96009  
 C/Species: *Sinorhizobium meliloti*  
 C/Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
 C/Accession: H96009  
 R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A/Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo  
 A/Reference number: A95842; MUID:21396508; PMID:11481431  
 A/Accession: H96009  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-275 <KUR>  
 A/Cross-references: GB:AL591985; PIDN:CA949744.1; PID:915141231; GSPDB:GN00167

A:Experimental source: strain 1021, megaplasmid pSMB  
 R:Gallibery, P.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Barloy-Hubler, P.; Chaitin, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002  
 C:Accession: AG3112  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woc, I.; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; PMID:11743193  
 A:Accession: AG3112  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-282 <KUR>  
 A:Cross-references: GB:AB008689; PIDN:AA145317.1; PID:G17743007; GSPDB:GN00187  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atu4523  
 A:Map position: linear chromosome  
 C:Superfamily: oligopeptide permease protein oppb

Query Match 0.7%; Score 7; DB 2; Length 275;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 362 EDLEGAV 368  
 DB 154 EDLEGAV 160

RESULT 41  
 AG3112  
 hypothetical protein Atu4523 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002  
 C:Accession: AG3112  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woc, I.; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; PMID:11743193  
 A:Accession: AG3112  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-282 <KUR>  
 A:Cross-references: GB:AB008689; PIDN:AA145317.1; PID:G17743007; GSPDB:GN00187  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atu4523  
 A:Map position: linear chromosome  
 C:Superfamily: oligopeptide permease protein oppb

Query Match 0.7%; Score 7; DB 2; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 920 ALLTTL 926  
 DB 255 ALLTTL 261

RESULT 42  
 AF2179  
 hypothetical protein all2989 [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp.  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
 C:Accession: AF2179  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Takata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AF2179  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-282 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA874688.1; PID:G17132083; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: all2989

Query Match 0.7%; Score 7; DB 2; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 752 KAGVSSQ 758  
 DB 204 KAGVSSQ 210

RESULT 43  
 F98174  
 oligopeptide ABC transporter (permease) (AP001508) [imported] - Agrobacterium tumefaciens  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002  
 C:Accession: F98174  
 R:Goodner, B.; Hinkle, G.; Gatung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.  
 Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
 A:Reference number: A97359; PMID:11743194  
 A:Accession: F98174  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-284 <KUR>  
 A:Cross-references: GB:AE007870; PIDN:AAK88920.1; PID:G15158694; GSPDB:GN00170  
 A:Genetics:  
 A:Gene: AGR L 692  
 A:Map position: linear chromosome  
 C:Superfamily: oligopeptide permease protein oppb

Query Match 0.7%; Score 7; DB 2; Length 284;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 920 ALLTTL 926  
 DB 257 ALLTTL 263

RESULT 44  
 A71065  
 probable sugar-binding transport system permease protein - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
 C:Accession: A71065  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki, M.; Ohtoku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi, M.  
 DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic bacterium Pyrococcus horikoshii  
 A:Reference number: A71000; MUID:98344137; PMID:9679194  
 A:Accession: A71065  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-292 <KAW>  
 A:Cross-references: GB:AP000005; NID:G3236132; PIDN:BA30315.1; PID:G3257632  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C:Genetics:  
 A:Gene: PH1215  
 C:Superfamily: inner membrane protein uppa

Query Match 0.7%; Score 7; DB 2; Length 292;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 826 VPGSLL 832  
 DB 79 VPGSLL 85

RESULT 45  
A87291  
succinyl-CoA synthetase, alpha subunit [imported] - *Caulobacter crescentus*  
C/Species: *Caulobacter crescentus*  
C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C/Accession: A87291  
R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kollar  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A/Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A/Reference number: A87249; MUID:21173698; PMID:11259647  
A/Accession: A87291  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-294 <STO>  
A/Cross-references: GB:AE005673; NID:g13421487; PIDN:AAK22325.1; GSPDB:GN00148  
C/Genetics:  
A/Gene: CC0338  
C/Superfamily: succinate-CoA ligase (ADP-forming) alpha chain

Query Match 0.7%; Score 7; DB 2; Length 294;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 DPKVQTE 90  
|||||  
DB 181 DPKVQTE 187

RESULT 46  
T22272  
hypothetical protein T01C4.3 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000  
C/Accession: T22272  
R/Graves, T.; Wohlmann, P.  
Submitted to the EMBL Data Library, September 1996  
A/Description: The sequence of C. elegans cosmid T01C4.  
A/Reference number: Z20599  
A/Accession: T22272  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-294 <GBA>  
A/Cross-references: EMBL:U70858; PIDN:AA09178.1; GSPDB:GN00023; CESP:T01C4.3  
A/Experimental source: strain Bristol N2; clone T01C4  
C/Genetics:  
A/Gene: CESP:T01C4.3  
A/Map position: 5  
A/Intons: 67/2, 116/2, 142/1, 220/3  
C/Superfamily: *Caenorhabditis elegans* hypothetical protein K07C6.10

Query Match 0.7%; Score 7; DB 2; Length 294;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 469 FMITLV 475  
|||||  
DB 39 FMITLV 45

RESULT 47  
E97675  
succinyl-CoA synthetase alpha chain (AP001515) [imported] - *Agrobacterium tumefaciens* (S  
C/Species: *Agrobacterium tumefaciens*  
C/Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C/Accession: E97675  
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*

A/Reference number: A97359; PMID:11743194  
A/Accession: E97675  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-300 <KUR>  
A/Cross-references: GB:AE007869; PIDN:AAK8358.1; PID:g15157841; GSPDB:GN00169  
C/Genetics:  
A/Gene: AGR\_C\_4779  
A/Map position: circular chromosome  
C/Superfamily: succinate-CoA ligase (ADP-forming) alpha chain

Query Match 0.7%; Score 7; DB 2; Length 300;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 DPKVQTE 90  
|||||  
DB 187 DPKVQTE 193

RESULT 48  
AD2900  
succinyl-CoA synthetase alpha chain [imported] - *Agrobacterium tumefaciens* (strain C58,  
C/Species: *Agrobacterium tumefaciens*  
C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002  
C/Accession: AD2900  
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, D.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McEl  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
A/Reference number: AB2577; PMID:11743193  
A/Accession: AD2900  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-300 <KUR>  
A/Cross-references: GB:AE008688; PIDN:AAL43618.1; PID:g17741139; GSPDB:GN00186  
A/Experimental source: strain C58 (Dupont)  
C/Genetics:  
A/Gene: snCD  
A/Map position: circular chromosome  
C/Superfamily: succinate-CoA ligase (ADP-forming) alpha chain

Query Match 0.7%; Score 7; DB 2; Length 300;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 DPKVQTE 90  
|||||  
DB 187 DPKVQTE 193

RESULT 49  
AF3269  
succinate-CoA ligase (ADP-forming) (BC 6.2.1.5) [imported] - *Brucella melitensis* (strain  
C/Species: *Brucella melitensis*  
C/Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 03-Jun-2002  
C/Accession: AF3269  
R/DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
.; Mazur, M.; Goldstein, E.; Selkov, E.; Elzer, P.H.; Haglue, S.; O'Callaghan, D.; Letes  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A/Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*  
A/Reference number: AD3252; PMID:11756688  
A/Accession: AF3269  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-300 <KUR>  
A/Cross-references: GB:AE008917; PIDN:AAL51321.1; PID:g17982017; GSPDB:GN00190  
A/Experimental source: strain 16M  
C/Genetics:  
A/Gene: BME10139

A:Map position: 1  
C:Superfamily: succinate-CoA ligase (ADP-forming) alpha chain  
C:Keywords: acid-thiol ligase; coenzyme A

Query Match 0.7%; Score 7; DB 2; Length 300;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 DPVKGTE 90  
DB 187 DPVKGTE 193

RESULT 50  
547764  
dipeptide transport system permease protein dpqc - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C>Date: 27-Jan-1995 #sequence revision 27-Jan-1995 #text\_change 01-Mar-2002  
C:Accession: S47764; S61433; A65153  
R:Plunkett, G.  
submitted to the EMBL Data Library, March 1994  
A:Reference number: S47666  
A:Accession: S47764  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-300 <PLU>  
A:Cross-references: EMBL:U00039; NID:G466582; PIDN:AMB18520.1; PID:G466681  
R:Abouhamad, W.N.; Manson, M.D.  
Mol. Microbiol. 14, 1077-1092, 1994  
A:Title: The dipeptide permease of Escherichia coli closely resembles other bacterial tr  
A:Reference number: S61431; MUID:95231288; PMID:7536291  
A:Accession: S61433  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-300 <ABO>  
A:Cross-references: EMBL:U08399; NID:G349224; PIDN:AAA23704.1; PID:G349227  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
A.; Rose, D.J.; Mau, B.; Sano, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: A65153  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-300 <BLAT>  
A:Cross-references: GB:AE000431; GB:U00096; NID:G1789957; PIDN:AAC76567.1; PID:G1789964;  
C:Genetics:  
A:Experimental source: strain K-12, substrain MG1655  
A:Gene: dpqc  
C:Superfamily: oligopeptide permease protein oppb

Query Match 0.7%; Score 7; DB 2; Length 300;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 920 ALLTVL 926  
DB 275 ALLTVL 281

RESULT 51  
F91181  
dipeptide transport system permease protein 2 EC64422 [imported] - Escherichia coli (str  
C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: F91181  
R:Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuwara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
A:Reference number: A9629; MUID:21156231; PMID:11258796  
A:Accession: F91181  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-300 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA837845.1; PID:G13363896; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain R1MD\_050952  
C:Genetics:  
A:Gene: EC64422  
C:Superfamily: oligopeptide permease protein oppb

Query Match 0.7%; Score 7; DB 2; Length 300;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 920 ALLTVL 926  
DB 275 ALLTVL 281

RESULT 52  
B86028  
dipeptide transport system permease protein 2 [imported] - Escherichia coli (strain O15  
C:Species: Escherichia coli  
C>Date: 16-Feb-2001 #sequence revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: B86028  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: B86028  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-300 <STO>  
A:Cross-references: GB:AE005174; NID:G12518266; PIDN:AAG58686.1; GSPDB:GN00145; UMGP:Z45  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: dpqc  
C:Superfamily: oligopeptide permease protein oppb

Query Match 0.7%; Score 7; DB 2; Length 300;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 920 ALLTVL 926  
DB 275 ALLTVL 281

RESULT 53  
A10486  
dipeptide transport system permease protein dpqc [imported] - Yersinia pestis (strain CC  
C:Species: Yersinia pestis  
C>Date: 02-Nov-2001 #sequence revision 02-Nov-2001 #text\_change 09-Nov-2001  
C:Accession: A10486  
R:Parhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: A10486  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-300 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC93461.1; PID:G15981907; GSPDB:GN00175  
C:Genetics:  
A:Gene: dpqc  
C:Superfamily: oligopeptide permease protein oppb

Query Match 0.7%; Score 7; DB 2; Length 300;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 920 ALLTVL 926

Db 275 ALLTVL 281

# RESULT 54

AG0983

C:Species: Salmonella enterica subsp. enterica serovar Typh  
C:Note: this species has also been called Salmonella typh  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AG0983

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dord, L.; White, N.; Farrar, S.; Moul, S.; O'Garra, P.  
Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A:Reference number: AB0502; PMID:11677608

A:Accession: AG0983

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1300 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD07995.1; PID:G16504980; GSPDB:GN00176

C:Genetics:

A:Gene: STY4170

C:Superfamily: oligopeptide permease protein oppB

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 300;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

# RESULT 55

E86591

C:Species: Chlamydia pneumoniae (strain J138)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001

C:Accession: E86591

R:Shital, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.

A:Reference number: A86491; PMID:10871362

A:Accession: E86591

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1303 <STO>

A:Cross-references: GB:BA000008; NID:G8979181; PIDN:BAA99015.1; GSPDB:GN00142

A:Experimental source: strain J138

C:Genetics:

A:Gene: CPJ0807

# Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 303;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

# RESULT 56

D72031

C:Species: Chlamydia pneumoniae (strain CWL029)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 05-May-2000

C:Accession: D72031

R:Kallman, S.; Mitchell, W.; Marathe, R.; Lämmler, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; PMID:99206606; PMID:10192388

A:Accession: D72031

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-303 <ARN>

A:Cross-references: GB:AB001662; GB:AE001363; NID:G4377118; PIDN:AAD18945.1; PID:G437711

A:Experimental source: strain CWL029

C:Genetics:

A:Gene: CPN0807

# Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 303;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

# RESULT 57

JC4347

C:Species: Vibrio anguillarum

C:Date: 21-Dec-1995 #sequence\_revision 08-Feb-1996 #text\_change 05-May-2000

C:Accession: JC4347

R:Milton, D.L.; Norqvist, A.; Wolf-Matz, H.

Gene 164, 95-100, 1995

A:Title: Sequence of a novel virulence-mediating gene, virc, from Vibrio anguillarum.

A:Reference number: JC4347; PMID:96060845; PMID:7590330

A:Accession: JC4347

A:Molecule type: DNA

A:Residues: 1-306 <MTL>

A:Cross-references: GB:U17054; NID:G576654; PIDN:AAA6984.1; PID:G576655

A:Experimental source: VAN20

C:Genetics:

A:Gene: cyag

C:Keywords: methyltransferase; S-adenosylmethionine; virulence

# Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 306;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

# RESULT 58

AB0078

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001

C:Accession: AB0078

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davis, R.M.; Davies, P.; Dougan, G.;  
Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; PMID:21470413; PMID:11563360

A:Accession: AB0078

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-310 <RUR>

A:Cross-references: GB:AL590842; PIDN:CAC89485.1; PID:G15978721; GSPDB:GN00175

C:Genetics:

A:Gene: YPO0631

C:Superfamily: conserved hypothetical protein H11364

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 310;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 HLSARV 15  
 |||||  
 Db 277 HLSARV 283

## RESULT 59

C97224  
 hypothetical protein CAC2635 [imported] - Clostridium acetobutylicum  
 C/Species: Clostridium acetobutylicum  
 C/Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 C/Accession: C97224  
 R/Molling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
 J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
 A/Reference number: A96900; MUID:21359325; PMID:21359325  
 A/Accession: C97224  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-315 <KUR>  
 A/Cross-references: GB:AE001437; PIDN:AAK80582.1; PID:G15025662; GSPDB:GN00168  
 A/Experimental source: Clostridium acetobutylicum ATCC824  
 C/Genetics:  
 A/Gene: CAC2635

Query Match 0.7%; Score 7; DB 2; Length 315;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 571 KIYSINV 577  
 |||||  
 Db 306 KIYSINV 312

## RESULT 60

J02217  
 peroxidase (EC 1.11.1.7) precursor, anionic - Japanese aspen x large-toothed aspen  
 C/Species: Populus sibirica x Populus grandidentata (Japanese aspen x large-toothed asp  
 C/Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 20-Jun-2000  
 C/Accession: J02217  
 R/Kawai, S.; Matsumoto, Y.; Kajita, S.; Yamada, K.; Katayama, Y.; Morohoshi, N.  
 Biosci. Biotechnol. Biochem. 57, 131-133, 1993  
 A/Title: Nucleotide sequence for the genomic DNA encoding an anionic peroxidase gene fr  
 A/Reference number: J02217; MUID:93163019; PMID:7764045  
 A/Accession: J02217  
 A/Molecule type: DNA  
 A/Residues: 1-318 <KAW>  
 A/Cross-references: GB:D11102; NID:G217996; PIDN:BA01877.1; PID:G217997  
 C/Genetics:  
 A/Introns: 68/3; 131/3; 186/1

C/Superfamily: peroxidase  
 C/Keywords: Chromoprotein, glycoprotein, heme, iron, metalloprotein, oxidoreductase  
 F/1-21/Domain: signal sequence #status predicted <SIG>  
 F/22-318/Product: peroxidase, anionic #status predicted <MAT>  
 F/32-111/Disulfide bonds: #status predicted  
 F/34-91,166,203,253/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F/59/active site: Arg #status predicted  
 F/63/108/Binding site: heme iron (His) (axial ligands) #status predicted  
 F/65-70/Disulfide bonds: #status predicted  
 F/117-314/Disulfide bonds: #status predicted  
 F/195-223/Disulfide bonds: #status predicted

Query Match 0.7%; Score 7; DB 2; Length 318;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 680 VTLAGP 686  
 |||||  
 Db 129 VTLAGP 135

## RESULT 61

S01774

stable inheritance 36K protein - plasmid NR1

C/Species: plasmid NR1  
 C/Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 05-Nov-1999  
 C/Accession: S01774  
 R/Taduchi, A.; Min, Y.N.; Kim, C.K.; Fan, Y.L.; Womble, D.D.; Rownd, R.H.  
 J. Mol. Biol. 202, 511-525, 1988  
 A/Title: Genetic organization and nucleotide sequence of the stability locus of IncFI  
 A/Reference number: S01774; MUID:89011976; PMID:3172224  
 A/Accession: S01774  
 A/Molecule type: DNA  
 A/Residues: 1-320 <TAB>  
 A/Cross-references: EMBL:X12777; NID:943001; PIDN:CAA31264.1; PID:943002  
 C/Genetics:  
 A/Gene: stbA  
 A/Gene: stbA  
 A/Gene: plasmid

Query Match 0.7%; Score 7; DB 2; Length 320;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 386 FFKTNS 392  
 |||||  
 Db 301 FFKTNS 307

## RESULT 62

A24920  
 36K para protein - Escherichia coli plasmid R1  
 C/Species: Escherichia coli  
 C/Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 18-Jun-1993  
 C/Accession: A24920  
 R/Gerdes, K.; Molin, S.  
 J. Mol. Biol. 190, 269-279, 1986  
 A/Title: Partitioning of plasmid R1. Structural and functional analysis of the para locu  
 A/Reference number: A24920  
 A/Accession: A24920  
 A/Molecule type: DNA  
 A/Residues: 1-320 <GER>  
 C/Genetics:  
 A/Gene: plasmid

Query Match 0.7%; Score 7; DB 2; Length 320;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 386 FFKTNS 392  
 |||||  
 Db 301 FFKTNS 307

## RESULT 63

F82904  
 conserved hypothetical U0330 [imported] - Ureaplasma urealyticum  
 C/Species: Ureaplasma urealyticum  
 C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C/Accession: F82904  
 R/Glaes, J.I.; Lefkowitz, E.J.; Glaes, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
 submitted to Genbank, February 2000  
 A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mir  
 A/Reference number: A82870  
 A/Accession: F82904  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-327 <GLA>  
 A/Cross-references: GB:AE002130; GB:AF222894; NID:G6899302; PIDN:AAF30739.1; GSPDB:GN001  
 A/Experimental source: serovar 3; biovar 1  
 C/Genetics:  
 A/Gene: U0330  
 A/Gene: code: SGC3

Query Match 0.7%; Score 7; DB 2; Length 327;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Qy 469 FMIITLV 475  
 |||||  
 Db 170 FMIITLV 176

# RESULT 64

A:Accession: A96997  
 A:Title: Ferriochrome transport permease CAC0768 (imported) - Clostridium acetobutylicum  
 C:Species: Clostridium acetobutylicum  
 C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 R:Accession: A96997  
 R:Author: J. J. Brelton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C10  
 A:Reference number: A96900; MUID:21359325; PMID:21359325  
 A:Accession: A96997  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-336 <NR>  
 A:Cross-references: GB:AE001437; PIDN:AAK76764.1; PID:g15023675; GSPDB:GN00168  
 A:Experimental source: Clostridium acetobutylicum ATCC824  
 A:Genetics: CAC0768

Query Match 0.7%; Score 7; DB 2; Length 336;  
 Best local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 921 ILITVLT 927  
 |||||  
 Db 19 ILITVLT 25

# RESULT 65

SS5491  
 probable galactosyltransferase (EC 2.4.1.-) SPAC5H10.13c [similarity] - fission yeast (S  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 01-Aug-1995 #sequence\_revision 01-Sep-1995 #text\_change 02-Mar-2001  
 C:Accession: T38977; SS5491  
 R:Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 submitted to the EMBL Data Library, May 1995  
 A:Reference number: Z21821  
 A:Accession: T38977  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-346 <CO2>  
 A:Cross-references: EMBL:Z49811; NID:9854599; PIDN:CAA89963.1; PID:9854612; GSPDB:GN0006  
 A:Experimental source: strain 972h; cosmid C5H10  
 C:Genetics:  
 A:Gene: SPAC5H10.13c  
 A:Map position: 1  
 C:Superfamily: Schizosaccharomyces pombe alpha-1,2-galactosyltransferase  
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 0.7%; Score 7; DB 2; Length 346;  
 Best local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 556 TTFHRS 562  
 |||||  
 Db 81 TTFHRS 87

# RESULT 66

A:DECHP  
 2-dehydro-3-deoxy-phosphonate aldolase (EC 4.1.2.15) (Phe-sensitive) - Escherichia  
 N:Alternate names: phospho-2-dehydro-3-deoxyheptonate aldolase; phospho-2-keto-3-deoxyhe  
 C:Species: Escherichia coli  
 C:Date: 18-Aug-1982 #sequence\_revision 17-Dec-1982 #text\_change 03-Jun-2002  
 C:Accession: A01106; B64811  
 R:Davies, W.D.; Davidson, B.E.

Nucleic Acids Res. 10, 4045-4058, 1982  
 A:Title: The nucleotide sequence of aroG, the gene for 3-deoxy-D-arabinohexulose-7  
 A:Reference number: A01106; MUID:82274236; PMID:6125934  
 A:Accession: A01106

A:Molecule type: DNA  
 A:Residues: 1-350 <DAV>  
 A:Cross-references: GB:U01591; NID:g145367; PIDN:AAA23492.1; PID:g145368  
 R:Experimental source: strain K12  
 A:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C  
 A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: B64811

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-350 <BLAT>

A:Cross-references: GB:AE000178; GB:U00096; NID:g1786967; PIDN:AACT3841.1; PID:g1786969

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: aroG

A:Map position: 17 min

C:Function:

A:Description: aldehyde-lyase; carbon-carbon lyase; this is one of the DAP synthases c  
 ic acid-7-phosphate

A:Pathway: aromatic amino acid biosynthesis; shikimate pathway

A:Note: the first reaction in aromatic amino acid biosynthesis

A:Note: feedback-inhibited by phenylalanine; the other two DAP synthases are tyrosine-  
 C:Superfamily: phospho-2-dehydro-3-deoxyheptonate aldolase

C:Keywords: aldehyde-lyase; aromatic amino acid biosynthesis; carbon-carbon lyase; shik

Query Match 0.7%; Score 7; DB 1; Length 350;  
 Best local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 AGEFLDM 103  
 |||||  
 Db 141 AGEFLDM 147

# RESULT 67

AG0593  
 phospho-2-dehydro-3-deoxyheptonate aldolase (DAP synthetase) phenylalanine repressible  
 C:Species: Salmonella enterica subsp. enterica serovar Typh  
 A:Note: this species has also been called Salmonella typh  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: AG0593  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
 th, T.; Connor, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
 S.; Mole, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.  
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero  
 A:Reference number: AB0502; PMID:11677608  
 A:Accession: AG0593

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-350 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD05217.1; PID:g16501987; GSPDB:GN00176

C:Genetics:

A:Gene: SRY0801

C:Superfamily: phospho-2-dehydro-3-deoxyheptonate aldolase

Query Match 0.7%; Score 7; DB 2; Length 350;  
 Best local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 AGEFLDM 103  
 |||||  
 Db 141 AGEFLDM 147

# RESULT 68

P90726  
 2-dehydro-3-deoxy-phosphoheptonate aldolase (EC 4.1.2.15) (Phe-sensitive) [similarity]  
 C:Species: Escherichia coli  
 C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Jun-2002  
 C:Accession: F90726  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
 gaawata, N.; Yasunaga, T.; Kubota, S.; Shiba, T.; Hatori, M.; Shingawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc  
 A:Reference number: A9629; PMID:1156231; PMID:1158796  
 A:Accession: F90726  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-350 <HAY>  
 A:Cross-references: GB:BA000007; PIDN:BAH34205.1; PID:g13360241; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RIMD 0509952  
 C:Genetics:  
 A:Gene: EC80782  
 C:Superfamily: phospho-2-dehydro-3-deoxyheptonate aldolase  
 C:Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 0.7%; Score 7; DB 2; Length 350;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AGEFLDM 103  
 |||||  
 DB 141 AGEFLDM 147

RESULT 69  
 2-dehydro-3-deoxy-phosphoheptonate aldolase (EC 4.1.2.15) (Phe-sensitive) [similarity]  
 C:Species: Escherichia coli  
 C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 03-Jun-2002  
 C:Accession: G85577  
 R:Penna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 11ler, U.; Grobeck, E.U.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Agodada,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; PMID:11206551  
 A:Accession: G85577  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-350 <STO>  
 A:Cross-references: GB:AE005174; NID:g12513683; PIDN:ANG55083.1; GSPDB:GN00145; UNCF:209  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: aroC  
 C:Superfamily: phospho-2-dehydro-3-deoxyheptonate aldolase  
 C:Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 0.7%; Score 7; DB 2; Length 350;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AGEFLDM 103  
 |||||  
 DB 141 AGEFLDM 147

RESULT 70  
 A97788  
 microcin C7 self-immunity protein homolog [imported] - Rickettsia conorii (strain Malish  
 C:Species: Rickettsia conorii  
 C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 30-Sep-2001  
 C:Accession: A97788  
 R:Ogata, H.; Audic, S.; Reneseo-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Rc  
 Science 293, 2033-2039, 2001  
 A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
 A:Reference number: A97700; PMID:11442074; PMID:11557893  
 A:Accession: A97788  
 A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-360 <RKR>  
 A:Cross-references: GB:AE006914; PIDN:AA102343.1; PID:g15619797; GSPDB:GN00173  
 C:Genetics:  
 A:Gene: mcf2

Query Match 0.7%; Score 7; DB 2; Length 360;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 874 GIOKTTY 880  
 |||||  
 DB 5 GIOKTTY 11

RESULT 71  
 164128  
 2-dehydro-3-deoxy-phosphoheptonate aldolase (EC 4.1.2.15) (Phe-sensitive) - Haemophilus  
 C:Species: Haemophilus influenzae  
 C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 03-Jun-2002  
 C:Accession: 164128  
 R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirnesh, E.F.; Kerlavage,  
 i Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.T.; Glodex, A.; Kelley, J.M.; Weidman,  
 D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
 Science 269, 496-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A:Reference number: A64000; PMID:95350630; PMID:7542800  
 A:Accession: 164128  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-362 <TIGR>  
 A:Cross-references: GB:U12830; GB:U42023; NID:g1574389; PIDN:AA23197.1; PID:g1574392;  
 C:Superfamily: phospho-2-dehydro-3-deoxyheptonate aldolase  
 C:Keywords: aldehyde-lyase; aromatic amino acid biosynthesis; carbon-carbon lyase

Query Match 0.7%; Score 7; DB 2; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AGEFLDM 103  
 |||||  
 DB 148 AGEFLDM 154

RESULT 72  
 F87100  
 L-alanine dehydrogenase [imported] - Mycobacterium leprae  
 C:Species: Mycobacterium leprae  
 C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
 C:Accession: F87100  
 R:Coile, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Hc  
 R.; Davies, R.M.; Devlin, K.; Duthey, S.; Fellwell, T.; Fraser, A.; Hamlin, N.; Holroyd,  
 eam, M.A.; Rutherford, K.W.  
 Nature 409, 1007-1011, 2001  
 A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq  
 A:Title: Massive gene decay in the leprosy bacillus.  
 A:Reference number: A86909; PMID:1128732; PMID:11234002  
 A:Accession: F87100  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-371 <STO>  
 A:Cross-references: GB:AL450380; NID:g13093357; PIDN:CA30483.1; GSPDB:GN00147  
 C:Genetics:  
 A:Gene: ald  
 C:Superfamily: alanine dehydrogenase; alanine dehydrogenase homology

Query Match 0.7%; Score 7; DB 2; Length 371;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 363 DLGAYK 369  
 |||||

Db 224 DLEGAVK 230

# RESULT 73

T20093

hypothetical protein C50B6.10 - *Caenorhabditis elegans*

C/Species: *Caenorhabditis elegans*

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 08-Dec-2000

C/Accession: T20093

R/Percy, C.

submitted to the EMBL Data Library, October 1996

A/Reference number: Z19222

A/Accession: T20093

A/Status: preliminary; translated from GB/EMBL/DDBT

A/Molecule type: DNA

A/Residues: 1-371 <Wtl>

A/Cross-references: EMBL:Z01050; PIDN:CA02859.1; GSPDB:GN00023; CESP:C50B6.10

A/Experimental source: clone C50B6

C/Genetics:

A/Map position: 5

A/Intons: 66/3; 96/1; 122/2; 167/2; 189/3; 210/3; 313/1

C/Superfamily: *Caenorhabditis elegans* hypothetical protein C45H4.12

## Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 371; Pred. No. 1.3e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 515 VGVSRT 521

Db 360 VGVSRT 366

# RESULT 74

A70791

probable anion transporting ATPase - *Mycobacterium tuberculosis* (strain H37RV)

C/Species: *Mycobacterium tuberculosis*

C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C/Accession: A70791

R/Colo, S.T.; Brosch, R.; Parthill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 399, 537-544, 1998

A/Authors: Saito, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.

A/Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome

A/Reference number: A70500; MUID:98295987; PMID:9634230

C/Accession: A70791

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-386 <COL>

A/Cross-references: GB:AL022121; GB:AL123456; NID:93261559; PIDN:CA118002.1; PID:el26454

A/Experimental source: strain H37RV

C/Genetics:

A/Gene: RV3680

## Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 386; Pred. No. 1.4e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 RLRLL 30

Db 178 RLRLL 184

# RESULT 75

T12456

hypothetical protein DKF2p564M2423.1 - human

C/Species: *Homo sapiens* (man)

C/Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 23-Jul-1999

C/Accession: T12456

R/Mambuti, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, June 1999

A/Reference number: Z17524

A/Accession: T12456

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-387 <MAN>

A/Cross-references: EMBL:AL080119

A/Experimental source: fetal brain; clone DKF2p564M2423

C/Genetics:

A/Note: DKF2p564M2423.1

## Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 387; Pred. No. 1.4e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 487 ADTENKE 493

Db 247 ADTENKE 253

# RESULT 76

C82637

porphyrin biosynthesis protein Xfr1797 [imported] - *Xylella fastidiosa* (strain 945c)

C/Species: *Xylella fastidiosa*

C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C/Accession: C82637

R/Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing

Nature 406, 151-157, 2000

A/Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A/Reference number: A82515; MUID:20365717; PMID:10910347

A/Note: for a complete list of authors see reference number A59328 below

A/Accession: C82637

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-389 <SIM>

A/Cross-references: GB:AE004001; GB:AE003849; NID:99106864; PIDN:AAF64605.1; GSPDB:GN001

A/Experimental source: strain 945c

R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Agencio, M.; Alvarenga, R.;

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrero, D.M.; Carre, R.

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to Genbank, June 2000

A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; From

J.D.; Unga, M.L.; Kemper, E.L.; Klatsjima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laiz

chado, M.A.; Madela, A.M.B.N.; Madela, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, R.

A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

Rodrigues, V.; Rosa, A.J. de M.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.F.

A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva

M.; Tashiro, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;

A/Reference number: A59328

A/Contents: annotation

C/Genetics:

A/Gene: Xfr1797

## Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 389; Pred. No. 1.4e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 988 LPPRGLL 994

Db 181 LPPRGLL 187

# RESULT 77

G89822

hypothetical protein SA0507 [imported] - *Staphylococcus aureus* (strain N315)

C/Species: *Staphylococcus aureus*

C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C/Accession: G89822

R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.;

ma, A.; Mizutani-U, T.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C/Species: *Staphylococcus aureus*

Lancet 357, 1225-1240, 2001

A/Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A/Reference number: A89758; MUID:21311952; PMID:11418146

A/Accession: G89822  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-391 <KUR>  
A/Cross-references: GB:BA000018; PID:g13700440; PIDN:BAB41738.1; GSPDB:GN00149  
A/Experimental source: strain N315  
C/Genetics:  
A/Gene: SA0507

Query Match 0.7%; Score 7; DB 2; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 TAILTV 925  
DB 108 TAILTV 114

RESULT 78  
AH0334  
probable long-chain fatty acid transport protein fadL [imported] - Versinia pestis (stra  
C/Species: Versinia pestis  
C/Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001  
C/Accession: AH0334  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tilball, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
Nature 413, 523-527, 2001  
A/Title: Genome sequence of Versinia pestis, the causative agent of plague.  
A/Reference number: AB0001; MUID:21470413; PMID:11586360  
A/Accession: AH0334  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-423 <KUR>  
A/Cross-references: GB:AL590842; PIDN:CAQ92983.1; PID:g15980722; GSPDB:GN00175  
C/Genetics:  
A/Gene: fadL  
C/Superfamily: long-chain fatty acid transport protein fadL

Query Match 0.7%; Score 7; DB 2; Length 423;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 824 KTVPGSL 830  
DB 256 KTVPGSL 262

RESULT 79  
AI2187  
histidinol dehydrogenase [imported] - Nostoc sp. (strain PCC 7120)  
C/Species: Nostoc sp.  
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C/Accession: AI2187  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iritani,  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A/Reference number: AB1807; MUID:21595285; PMID:11759840  
A/Accession: AI2187  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-431 <KUR>  
A/Cross-references: GB:BA000019; PIDN:BAB74755.1; PID:g17132150; GSPDB:GN00179  
A/Experimental source: strain PCC 7120  
C/Genetics:  
A/Gene: hisD  
C/Superfamily: histidinol dehydrogenase; histidinol dehydrogenase homology

Query Match 0.7%; Score 7; DB 2; Length 431;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 HSHLSA 12  
DB 61 HSHLSA 67

RESULT 80  
H90131  
regulator of epidermal growth factor receptor [imported] - Guillardia theta nucleomorph  
C/Species: nucleomorph Guillardia theta  
A/Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C/Accession: H90131  
R:Douglas, S.; Zanner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re  
Nature 410, 1091-1096, 2001  
A/Title: The highly reduced genome of an enslaved algal nucleus.  
A/Reference number: A99082; MUID:11323671; PMID:11323671  
A/Accession: H90131  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-434 <DOU>  
A/Cross-references: GB:AF083031; NID:g13794405; PIDN:AAK39782.1; GSPDB:GN00152  
C/Genetics:  
A/Gene: Ebi  
A/Map position: 3  
A/Genome: nucleomorph  
C/Keywords: nucleomorph

Query Match 0.7%; Score 7; DB 2; Length 434;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 933 KNOXLEY 939  
DB 304 KNOXLEY 310

RESULT 81  
T36706  
hypothetical protein SCH69.07c - Streptomyces coelicolor  
C/Species: Streptomyces coelicolor  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C/Accession: T36706  
R:Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrett, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, June 1999  
A/Reference number: Z21612  
A/Accession: T36706  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-436 <MUR>  
A/Cross-references: EMBL:AL079308; PIDN:CA845204.1; GSPDB:GN00070; SCOEDB:SCH69.07c  
A/Experimental source: strain A3(2)  
C/Genetics:  
A/Gene: SCOEDB:SCH69.07c

Query Match 0.7%; Score 7; DB 2; Length 436;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 679 TVTLAAG 685  
DB 187 TVTLAAG 193

RESULT 82  
J01729  
ankyrin-repeat protein - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 22-Jun-1999  
C/Accession: J01729  
R:Zhang, H.; Schelker, D.C.; Fowle, W.H.; Goodman, H.M.  
Plant Cell 4, 1575-1588, 1992  
A/Title: Expression of antisense or sense RNA of an ankyrin repeat-containing gene block

A;Reference number: JQ1729; MUID:93104681; PMID:1281700  
 A;Accession: JQ1729  
 A;Molecule type: mRNA  
 A;Residues: 1-439 <ZNA>  
 A;Cross-references: GB:M82803; NID:g16743; PIDN:AAA32812.1; PID:g16744  
 C;Genetics:  
 A;Gene: AKR  
 A;Intons: 262/1; 287/3; 320/3; 353/3; 386/3; 419/3  
 C;Superfamily: unassigned ankryrin repeat proteins; ankryrin repeat homology; BGF homology  
 C;Keywords: cytoskeleton  
 F;282-314/Domain: ankryrin repeat <RP1>  
 F;315-347/Domain: ankryrin repeat <RP2>  
 F;348-380/Domain: ankryrin repeat <RP3>  
 F;381-413/Domain: ankryrin repeat <RP4>  
 F;387-419/Domain: ankryrin repeat homology <AN1>

Query Match  
 Best Local Similarity 100.0%; Score 7; DB 2; Length 439;  
 Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 829 SLIPGT 835  
 DB 40 SLIPGT 46

RESULT 83  
 A44751  
 carotenoid-binding protein cbpa - *Synechococcus* sp.  
 C;Species: *Synechococcus* sp.  
 C;Date: 03-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 08-Oct-1999  
 C;Accession: A44751  
 R;Reddy, K.J.; Masamoto, K.; Sherman, D.M.; Sherman, L.A.  
 J. Bacteriol. 171, 3486-3493, 1989  
 A;Title: DNA sequence and regulation of the gene (cbpa) encoding the 42-kilodalton cytop  
 A;Reference number: A44751; MUID:89255122; PMID:2498292  
 A;Accession: A44751  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-450 <RED>  
 A;Cross-references: EMBL:M27055; NID:g154493; PIDN:AAA27303.1; PID:g154494  
 A;Note: the authors translated the codon ATC for residue 119 as Met

Query Match  
 Best Local Similarity 100.0%; Score 7; DB 2; Length 450;  
 Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 460 TAAGASD 466  
 DB 47 TAAGASD 53

RESULT 84  
 S45051  
 tubulin alpha-B chain - *Neurospora crassa*  
 C;Species: *Neurospora crassa*  
 C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Aug-1999  
 C;Accession: S45051  
 R;Womac, J.; Turian, G.  
 Submitted to the EMBL Data Library, May 1994  
 A;Description: Two alpha-tubulin genes of *Neurospora crassa* encode divergent proteins.  
 A;Reference number: S45050  
 A;Accession: S45051  
 A;Molecule type: mRNA  
 A;Residues: 1-451 <MON>  
 A;Cross-references: EMBL:X79404; NID:g495132; PIDN:CAA55941.1; PID:g495133  
 C;Superfamily: tubulin.  
 C;Keywords: microtubule

Query Match  
 Best Local Similarity 100.0%; Score 7; DB 2; Length 451;  
 Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 964 MEGDVE 970

DB 442 MEGDVE 448

RESULT 85  
 T47824  
 hypothetical protein F24G16.270 - *Arabidopsis thaliana*  
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 28-Jul-2000  
 C;Accession: T47824  
 R;D'Angelio, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.  
 Submitted to the Protein Sequence Database, February 2000  
 A;Reference number: Z24477  
 A;Accession: T47824  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-451 <DNA>  
 A;Cross-references: EMBL:AL138647  
 A;Experimental source: cultivar Columbia; BAC clone F24G16  
 C;Genetics:  
 A;Map position: 3  
 A;Intons: 284/3; 324/3; 363/3; 393/3; 429/3  
 A;Note: F24G16.270  
 C;Superfamily: *Arabidopsis thaliana* hypothetical protein F24G16.270

Query Match  
 Best Local Similarity 100.0%; Score 7; DB 2; Length 451;  
 Pred. No. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 321 SSGCNVR 327  
 DB 210 SSGCNVR 216

RESULT 86  
 AB1104  
 DNA repair protein Sms homolog lmo023 [imported] - *Listeria monocytogenes* (strain EGD-e  
 C;Species: *Listeria monocytogenes*  
 C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
 C;Accession: AB1104  
 R;Glasner, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauget, O.; Ertlan, K.D.; Fsihi, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A;Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M  
 ok, C.; Schlietner, T.; Simoes, N.; Tlaxer, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund  
 A;Title: Comparative genomics of *Listeria* species.  
 A;Reference number: AB1077; MUID:21537279; PMID:11679669  
 A;Accession: AB1104  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-457 <GLA>  
 A;Cross-references: GB:NC\_003210; PIDN:CAD00760.1; PID:g16409598; GSPDB:GN00177  
 A;Experimental source: strain EGD-e  
 C;Genetics:  
 A;Gene: lmo023  
 C;Superfamily: DNA repair protein sms

Query Match  
 Best Local Similarity 100.0%; Score 7; DB 2; Length 457;  
 Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 ECTATLM 178  
 DB 195 ECTATLM 201

RESULT 87  
 T33787  
 hypothetical protein F52F10.2 - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 20-Jun-2000  
 C;Accession: T33787

R:Graves, T.; Henkhaus, J.; Wohldmann, J.; Bauer, C.; Duckels, G.  
A:Submitted to the EMBL Data Library, October 1998  
A:Description: The sequence of C. elegans cosmid F52F10.  
A:Reference number: Z21410  
A:Accession: T33787  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-471 <GRA>  
A:Cross-references: EMBL:AF101316; PIDN:AC69232.1; GSPDB:GN00023; CESP:F52F10.2  
A:Experimental source: strain Bristol N2; clone F52F10  
C:Genetics:  
A:Gene: CESP:F52F10.2  
A:Map position: 5  
A:Insertions: 39/1; 122/2; 203/1; 261/1; 308/2; 355/2; 448/3  
C:superfamily: Caenorhabditis elegans hypothetical protein C25E10.3

Query Match  
Best Local Similarity 100.0%; Score 7; DB 2; Length 471;  
Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

311 CDDPKYS 317  
397 CDDPKYS 403

RESULT 88  
E84853  
hypothetical protein At2g42400 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: E84853  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanham, L.; Tallon, L.;  
Eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: E84853  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-473 <STO>  
A:Cross-references: GB:AE002093; NID:94567312; PIDN:AAD33723.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g42400  
A:Map position: 2

Query Match  
Best Local Similarity 100.0%; Score 7; DB 2; Length 473;  
Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

782 AELFMLE 788  
335 AELFMLE 341

RESULT 89  
A64472  
carbamoyl-phosphate synthase, large chain (EC 6.3.-.-) - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A64472  
R:Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
; Reich, C.I.; Overbeek, R.; Kirschner, S.F.; Weissbach, K.G.; Merrick, J.M.; Glodek, A.;  
Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1085-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999; PMID:8688087  
A:Accession: A64472  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-487 <BUL>  
A:Cross-references: GB:U67577; GB:L77117; NID:91592013; PID:91592023; TIGR:MTJ378; PID:9

C:Genetics:  
A:Map position: FOR1326866-1328329  
A:Start codon: TTG  
C:Superfamily: carbamoyl-phosphate synthase large chain; biotin carboxylase homology  
C:Keywords: ligase  
F:11-470/Domain: biotin carboxylase homology <BCH>

Query Match  
Best Local Similarity 100.0%; Score 7; DB 1; Length 487;  
Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

689 TSKGLKY 695  
194 TSKGLKY 200

RESULT 90  
AF1130  
transcription regulator homolog lmo0445 [imported] - Listeria monocytogenes (strain EGD-  
C:Species: Listeria monocytogenes  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AF1130  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,  
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fsihl, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; Me  
ok, C.; Schlueber, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AF1130  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-495 <GUA>  
A:Cross-references: GB:NC\_003210; PIDN:CAC98524.1; PID:916409822; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo0445

Query Match  
Best Local Similarity 100.0%; Score 7; DB 2; Length 495;  
Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

490 ENKEVAR 496  
278 ENKEVAR 284

RESULT 91  
T01937  
hypothetical protein F1104.7 - Arabidopsis thaliana (fragment)  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 24-Mar-1999  
C:Accession: T01937  
R:Abu-Threideh, J.; Stoneking, T.; Langston, Y.; Trevasakis, E.  
submitted to the EMBL Data Library, October 1998  
A:Description: The sequence of A. thaliana F1104.  
A:Reference number: Z14466  
A:Accession: T01937  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-508 <ABU>  
A:Cross-references: EMBL:AF096370; NID:93695372; PID:93695380  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Map position: 4  
A:Note: F1104.7

Query Match  
Best Local Similarity 100.0%; Score 7; DB 2; Length 508;  
Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

364 LEGAVKL 370

Db 445 LEGAVKL 451

## RESULT 92

penton base protein - fowl adenovirus 1

N/Alterate names: envelope protein

C/Species: Aviadonovirus gall (fowl adenovirus 1, CEL0)

C/Date: 08-May-1995 #sequence\_revision 14-Feb-1997 #text\_change 26-Aug-1999

C/Accession: S52321

R/Akopian, T.A.; Lazareva, S.E.; Tikhomirov, E.E.; Karpov, V.A.; Narditsky, B.S. submitted to the EMBL Data Library, February 1995

A/Description: Genes for Avian adenovirus CEL0 penton base and core polypeptides.

A/Reference number: S52320

A/Accession: S52321

A/Molecule type: DNA

A/Residues: 1-515 <AKO>

A/Cross-references: EMBL:Z48167; NID:g755698; PIDN:CAA88181.1; PID:g663165

C/Superfamily: adenovirus penton protein

C/Keywords: envelope protein

## Query Match

Best Local Similarity 100.0%; Pred. No. 1.7e+02; Length 515;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 720 DLRIPEG 726

Db 181 DLRIPEG 187

## RESULT 93

nucleolar protein - human

C/Species: Homo sapiens (man)

C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000

C/Accession: J03035

R/Duki, N.; Kondo, M.; Seki, N.; Yano, K.; Oda, T.; Maehuo, Y.; Muramatsu, M. Biochem. Biophys. Res. Commun. 252: 97-102, 1998

A/Title: NOLP: Identification of a novel human nucleolar protein and determination of se

A/Reference number: J03035; PMID:99032811; PMID:9813152

A/Accession: J03035

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-524 <DEK>

A/Cross-references: DDBJ:AB017800; NID:g3892915; PIDN:BAA34576.1; PID:g3892916

## Query Match

Best Local Similarity 100.0%; Pred. No. 1.7e+02; Length 524;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 DSDAAS 150

Db 144 DSDAAS 150

## RESULT 94

high affinity nitrate transporter protein 2 AtNRT2-1 [similarity] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 02-Sep-2000

C/Accession: T51836

R/Fillier, S.; Daniel-Vedele, F. Planta 207: 461-469, 1999

A/Title: Expression analysis of a high affinity nitrate transporter from Arabidopsis th

A/Reference number: 225479; PMID:99136914; PMID:9951738

A/Accession: T51836

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: mRNA

A/Residues: 1-530 <FIL>

A/Cross-references: EMBL:AF093754; PIDN:AAC64170.1

A/Experimental source: cultivar Landsberg erecta

C/Genetics:

A/Map position: 1  
C/Superfamily: nitrate transporter component  
C/Keywords: nitrate transport

## Query Match

Best Local Similarity 100.0%; Pred. No. 1.8e+02; Length 530;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 DPKVGT 90

Db 474 DPKVGT 480

## RESULT 95

hypothetical protein lctp [imported] - Staphylococcus aureus (strain N315)

C/Species: Staphylococcus aureus

C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C/Accession: B89771

R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaishi, H.; Hiramatsu, K. Lancet 357: 1225-1240, 2001

A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A/Reference number: B89758; PMID:21311952; PMID:11418146

A/Accession: B89771

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-530 <KUR>

A/Cross-references: GB:BA000018; PID:g13700026; PIDN:BA841325.1; GSPDB:GN00149

A/Experimental source: strain N315

C/Genetics:

C/Superfamily: L-lactate permease

## Query Match

Best Local Similarity 100.0%; Pred. No. 1.8e+02; Length 530;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 TAILTLY 925

Db 358 TAILTLY 364

## RESULT 96

Ro/SS-A complex, 60K ribonucleoprotein - human

C/Species: Homo sapiens (man)

C/Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 22-Oct-1999

C/Accession: A31760; A30596

R/Deutscher, S.L.; Harley, J.B.; Keene, J.D. Proc. Natl. Acad. Sci. U.S.A. 85: 9479-9483, 1988

A/Title: Molecular analysis of the 60-kDa human Ro ribonucleoprotein.

A/Reference number: A31760; PMID:89071722; PMID:3200833

A/Accession: A31760

A/Molecule type: mRNA

A/Residues: 1-538 <DEU>

A/Cross-references: GB:U04137; NID:g177782; PIDN:AAA35493.1; PID:g177783

A/Ben-Cherif, E.; Gandy, B.J.; Tan, E.M.; Sullivan, K.F. J. Clin. Invest. 83: 1284-1292, 1989

A/Title: Isolation and characterization of a cDNA clone encoding the 60-kD component of

A/Reference number: A30596; PMID:8918084; PMID:2649513

A/Accession: A30596

A/Molecule type: mRNA

A/Residues: 1-238; 'R', 240-292, 'DV', 295-300, 'A', 302-514, 'ALQNTLNKSR' <BEN>

A/Cross-references: GB:M25077; NID:g387656

A/Note: the sequence is revised in GenBank entry HUMANTANP, release 111.0, (PID:g387656)

C/Genetics:

A/Gene: GDB:SSA2

A/Cross-references: GDB:355563; OMIM:600063

A/Map position: 1q31-1q31

C/Keywords: DNA binding; zinc finger

F;305-323/Region: zinc finger CCH motif



Query Match 0.7%; Score 7; DB 2; Length 538;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 769 VTTDMTL 775  
 |||||  
 DB 424 VTTDMTL 430

RESULT 97  
 T36704  
 probable dehydrogenase - Streptomyces coelicolor  
 C/Species: Streptomyces coelicolor  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C/Accession: T36704  
 R/Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, June 1999  
 A/Reference number: Z21612  
 A/Accession: T36704  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-561 <MUR>  
 A/Cross-references: EMBL:AL079308; PIDN:CAB45202.1; GSPDB:GN00070; SCOEDB:SCH69.05  
 A/Experimental source: strain A3(2)  
 C/Genetics:

A/Genes: SCOEDB:SCH69.05

Query Match 0.7%; Score 7; DB 2; Length 563;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 270 AYTSECF 276  
 |||||  
 DB 451 AYTSECF 457

RESULT 98  
 A30098  
 developmental protein xlgv7 - African clawed frog  
 C/Species: Xenopus laevis (African clawed frog)  
 C/Date: 21-May-1990 #sequence\_revision 21-May-1990 #text\_change 16-Feb-1997  
 C/Accession: A30098  
 R/Miller, M.; Kloc, M.; Reddy, B.; Eastman, E.; Dreyer, C.; Eskin, L.  
 Genes Dev. 3, 572-583, 1989  
 A/Title: xlgv7: a maternal gene product localized in nuclei of the central nervous system  
 A/Reference number: A30098; MUID:89252831; PMID:2721962  
 A/Accession: A30098  
 A/Molecule type: mRNA  
 A/Residues: 1-591 <MIL>  
 A/Cross-references: EMBL:J04494  
 C/Keywords: DNA binding

Query Match 0.7%; Score 7; DB 2; Length 591;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 739 CQAVIIP 745  
 |||||  
 DB 509 CQAVIIP 515

RESULT 99  
 T38920  
 methyltetrahydrofolate reductase 2 - fission yeast (Schizosaccharomyces pombe)  
 C/Species: Schizosaccharomyces pombe  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C/Accession: T38920  
 R/Pearson, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 submitted to the EMBL Data Library, February 1996  
 A/Reference number: Z21817  
 A/Accession: T38920  
 A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA  
 A/Residues: 1-603 <PEA>  
 A/Cross-references: EMBL:Z69728; PIDN:CAA93581.1; GSPDB:GN00066; SPDB:SPAC56F8.10  
 A/Experimental source: strain 972h-; cosmid c56F8  
 C/Genetics:  
 A/Genes: SPDB:SPAC56F8.10  
 A/Map position: 1

Query Match 0.7%; Score 7; DB 2; Length 603;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 WDELPHG 133  
 |||||  
 DB 342 WDELPHG 348

RESULT 100  
 T38908  
 probable gamma-glutamyltransferase (EC 2.3.2.2) SPAC56E4.06c, precursor [similarity] -  
 C/Species: Schizosaccharomyces pombe  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 02-Jun-2000  
 C/Accession: T38908  
 R/Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
 submitted to the EMBL Data Library, September 1997  
 A/Reference number: Z21813  
 A/Accession: T38908  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-611 <CON>  
 A/Cross-references: EMBL:Z99261; PIDN:CAB16397.1; GSPDB:GN00066; SPDB:SPAC56E4.06c  
 A/Experimental source: strain 972h-; cosmid c56E4  
 C/Genetics:  
 A/Genes: SPDB:SPAC56E4.06c  
 A/Map position: 1  
 C/Superfamily: gamma-glutamyltransferase  
 C/Keywords: aminoacyltransferase

Query Match 0.7%; Score 7; DB 2; Length 611;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 727 ESGFSKS 733  
 |||||  
 DB 555 ESGFSKS 561

Search completed: April 22, 2003, 15:36:05  
 Job time : 50 secs

